

# UNCLASSIFIED

AD NUMBER
AD011390
NEW LIMITATION CHANGE
TO Approved for public release, distribution unlimited
FROM Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; 01 MAY 1953. Other requests shall be referred to Office of Naval Research, Arlington, VA 22203.
AUTHORITY
ST-A per ONR ltr dtd 9 Nov 1977

THIS PAGE IS UNCLASSIFIED

Reproduced by  
**Armed Services Technical Information Agency**  
**DOCUMENT SERVICE CENTER**  
KNOTT BUILDING, DAYTON, 2, OHIO

**AD -**

**11390**

**UNCLASSIFIED**

11390

# TRANSFORMATIONS OF SYSTEMS OF RELATIVISTIC PARTICLE MECHANICS

BY  
HERMAN RUBIN and PATRICK SUPPES

TECHNICAL REPORT NO. 2  
MAY 1, 1953

PREPARED UNDER CONTRACT Nonr 225 (13)  
(NR-045-120)  
FOR  
OFFICE OF NAVAL RESEARCH

DEPARTMENT OF PHILOSOPHY  
AND  
APPLIED MATHEMATICS AND STATISTICS LABORATORY  
STANFORD UNIVERSITY  
STANFORD, CALIFORNIA

TRANSFORMATIONS OF SYSTEMS OF RELATIVISTIC  
PARTICLE MECHANICS

BY  
HERMAN RUBIN AND PATRICK SUPPES

TECHNICAL REPORT NO. 2

MAY 1, 1953

PREPARED UNDER CONTRACT Nonr-225(13)  
(NR-045-120)

FOR  
OFFICE OF NAVAL RESEARCH

DEPARTMENT OF PHILOSOPHY  
AND  
APPLIED MATHEMATICS AND STATISTICS LABORATORY  
STANFORD UNIVERSITY  
STANFORD, CALIFORNIA

TRANSFORMATIONS OF SYSTEMS OF RELATIVISTIC  
PARTICLE MECHANICS<sup>1/</sup>

by

Herman Rubin and Patrick Suppes

1 Introduction.

In McKinsey, Sugar, and Suppes [7] the axiomatic foundations of classical particle mechanics were investigated; and in McKinsey and Suppes [8] the transformations which carry systems of classical particle mechanics into systems of classical particle mechanics were determined. The purpose of the present paper is a similar investigation of relativistic particle mechanics (in the sense of the special theory of relativity). Some remarks on the general orientation of these studies are to be found in Section 1 of McKinsey, Sugar, and Suppes [7], and in McKinsey and Suppes [9].

In regard to our axiomatization of relativistic particle mechanics, we want to emphasize that we have in no sense attempted to use primitive notions which are logically or epistemologically simple. Investigations with these latter aims are to be found in Reichenbach [11], Robb [12], Schnell [13], and Walker [14]; but these studies are incomplete in the sense that they do not give axioms adequate for relativistic particle mechanics as it is ordinarily conceived by physicists. We have attempted to present such a complete set of axioms in a mathematically clear way.

The main result of the present paper is the determination under a certain weak hypothesis of the set of transformations which always carry systems of relativistic particle mechanics into systems of relativistic particle mechanics. Although this set of transformations is not a group (under the usual operation)

we are able to show that it is essentially a Brandt groupoid. It is difficult precisely to compare our results with those of MacColl [6], but our results seem to represent an improvement in three respects: (i) we work within an explicit axiomatic framework; (ii) we consider transformations of the units of mass and force as well as position and time; (iii) we consider transformations from one value for the velocity of light to another.

We briefly summarize the mathematical notations we use, most of which are standard. We denote the ordered  $n$ -tuple whose first member is  $a_1$ , whose second member is  $a_2$ , and so on, by

$$\langle a_1, \dots, a_n \rangle.$$

By an  $n$ -dimensional vector we mean an ordered  $n$ -tuple of real numbers.

Operations on vectors are defined in the usual way. We use the symbol "0" to denote the real number zero, the  $n$ -dimensional vector all of whose components are zero, and the matrix all of whose elements are zero. If  $A = \langle a_1, \dots, a_n \rangle$  is any vector, the length  $|A|$  of  $A$  is defined by

$$|A| = \sqrt{a_1^2 + \dots + a_n^2},$$

and by  $[A]_{i_1, i_2, \dots, i_r}$ , we mean the  $r$ -dimensional vector  $\langle a_{i_1}, a_{i_2}, \dots, a_{i_r} \rangle$ .

Thus if  $A = \langle 4, 7, 5 \rangle$ , then  $[A]_{2,3} = \langle 7, 5 \rangle$ . If  $A$  is a vector, we sometimes write " $A^2$ " for " $|A|^2$ ". If  $A$  is a matrix, we denote the transpose of  $A$  by " $A^*$ ", and the determinant of  $A$  by " $|A|$ ". We denote the identity matrix by " $I$ ". Although we treat vectors as one-rowed matrices, if  $A$  is a vector we always mean by  $|A|$ , the length of  $A$  and not the determinant of  $A$ : the meaning should be clear from the context. We use both matrix notation and usual vector notation for the inner product of two vectors  $A$  and  $B$ . Thus, we sometimes write:  $AB^*$ , and sometimes:  $A \cdot B$ , whichever is more convenient.

We use Menger's notation for derivatives (see Menger [10]). If  $f$  is a function, then  $D(f)$  is the derivative of  $f$ . Thus, for example,  $D(\sin) = \cos$ ,  $[D(\sin)](x) = \cos x$ , and  $[D^2(\sin)](x) = -\sin x$ . In this connection, we use the standard notation for sums, products, quotients, square roots, etc. of functions. Thus, for example, if  $f$  and  $g$  are functions of a real variable, by  $f+g$  we mean the function  $h$  such that for every real number  $x$

$$h(x) = f(x) + g(x)$$

If  $f$  is a 1-1 function,  $f^{-1}$  is the inverse function of  $f$ . It is also convenient to introduce a special symbol for the composition of two functions: if  $f$  and  $g$  are functions of a real variable, by  $g \circ f$  we mean the function  $h$  such that for every real number  $x$

$$h(x) = g(f(x))$$

To make some of our equations involving derivatives more perspicuous in relation to the notation ordinarily used in physics, we introduce formally the following two symbols: if  $f$  and  $g$  are functions of a real variable, then the function  $\frac{df}{dg}$  is defined by the following equation (for all real numbers  $x$ )

$$\frac{df}{dg}(x) = \left( \frac{Df}{Dg} \right)(x),$$

and the function  $\frac{d^2f}{dg^2}$  by the equation

$$\frac{d^2f}{dg^2}(x) = \left( \frac{D\left[\frac{Df}{Dg}\right]}{Dg} \right)(x).$$

Finally, we also use the following notation:  $I$  is the set of all positive integers,  $R$  is the set of all real numbers,  $R^+$  is the set of all positive real numbers, and  $E_n$  is the set of all  $n$ -dimensional vectors. We sometimes use geometrical language, referring to vectors in  $E_n$  as points in  $n$ -dimensional Euclidean space, etc.

## 2. Primitive notions.

Our axioms for relativistic particle mechanics are based on six primitive notions:  $P$ ,  $\mathcal{A}$ ,  $m$ ,  $s$ ,  $f$ , and  $c$ .  $P$  is a set,  $\mathcal{A}$  and  $m$  are unary functions,  $s$  is a binary function,  $f$  is a ternary function, and  $c$  is a constant.

The intended physical interpretation of  $P$  is as the set of particles. For every  $p$  in  $P$ ,  $\mathcal{A}(p)$  is to be interpreted physically as a set of real numbers measuring elapsed times (in terms of some unit of time and measured from some origin of time). There is a good physical reason for assigning (possibly) different sets of real numbers to different particles, instead of having one set of elapsed times for the whole system, as in McKinsey, Sugar, and Suppes [7]: two particles which have a simultaneous "life-span" with respect to one inertial frame of reference may have life-spans which do not even overlap with respect to another inertial frame.

For every  $p$  in  $P$ ,  $m(p)$  is to be interpreted physically as the numerical value of the rest mass of  $p$ . For every  $p$  in  $P$  and  $t$  in  $\mathcal{A}(p)$ ,  $s(p,t)$  is a vector, to be thought of physically as giving the position of  $p$  at time  $t$ . Thus the primitive  $s$  fixes the choice of a coordinate system. It is also possible to take as a primitive the set of all admissible (i.e., inertial) coordinate systems; this procedure is followed in Hermes [3]. We remark that for a fixed  $p$  in  $P$ , it is usually convenient to use in place of  $s$  the function  $s_p$ , which is defined on  $\mathcal{A}(p)$  and is such that for every  $t$  in  $\mathcal{A}(p)$ ,  $s_p(t) = s(p,t)$ .

For every  $p$  in  $P$  and  $t$  in  $\mathcal{A}(p)$ , and for  $i$  any positive integer,  $f(p,t,i)$  is a vector giving the components (parallel to the axes of the coordinate system) of the  $i^{\text{th}}$  force acting on  $p$  at time  $t$ . For further discussion of this primitive, applicable to relativistic as well as classical particle mechanics, see McKinsey, Sugar, and Suppes [7].

Our primitive constant  $c$  is to be interpreted as the numerical value of the velocity of light.



### 3. Axioms.

Using the six primitive notions just described, we now give our axioms for relativistic particle mechanics.

An ordered sextuple  $\Gamma = \langle P, \mathcal{T}, m, s, f, c \rangle$  which satisfies Axioms A1-A7 is called an n-DIMENSIONAL SYSTEM OF RELATIVISTIC PARTICLE MECHANICS (or simply a SYSTEM OF RELATIVISTIC PARTICLE MECHANICS, for abbreviation, S.R.

#### Kinematical axioms

- A1. P is a non-empty, finite set.
- A2. If  $p \in P$ , then  $\mathcal{T}(p)$  is an interval of real numbers.
- A3. If  $p \in P$  and  $t \in \mathcal{T}(p)$ , then  $s_p(t)$  is an n-dimensional vector; and moreover, the second derivative of  $s_p$  exists throughout the interval  $\mathcal{T}(p)$ .
- A4. The constant  $c$  is a positive real number such that for every  $p \in P$  and  $t$  in  $\mathcal{T}(p)$ ,  

$$|(Ds_p)(t)| < c.$$

#### Dynamical axioms

- A5. If  $p \in P$ , then  $m(p)$  is a positive real number.
- A6. If  $p \in P$  and  $t \in \mathcal{T}(p)$ , then  $f(p, t, 1), f(p, t, 2), \dots$  are n-dimensional vectors such that the series

$$\sum_{i=1}^{\infty} f(p, t, i)$$

is absolutely convergent.

- A7. If  $p \in P$  and  $t \in \mathcal{T}(p)$ , then

$$m(p) \left[ D \left( \frac{Ds_p}{\sqrt{1 - \frac{|Ds_p|^2}{c^2}}} \right) \right](t) = \sqrt{1 - \frac{|(Ds_p)(t)|^2}{c^2}} \cdot \sum_{i=1}^{\infty} f(p, t, i)$$

Since this set of axioms is similar in many ways to that given for classical mechanics in McKinsey, Sugar, and Suppes [7], a large number of remarks to be found in Section 3 of that paper are also applicable here and will not be repeated. From Axiom A7 it is clear that the force concept we are using is that of Minkowski. In the solution of special problems this concept is not always the most useful one, but the relative simplicity of its transformation properties more than justifies its use here. Some readers may feel that there are good physical grounds for taking the notion of relativistic mass as primitive instead of that of rest mass; however, it is easy to define the notion of relativistic mass in terms of the notion of rest mass and our other primitives, and the use of the notion of rest mass as a primitive emphasizes the considerable formal similarity between our axioms for relativistic mechanics and the axioms for classical mechanics of McKinsey, Sugar, and Suppes.

For  $p$  in  $P$ ,  $\mathcal{I}(p)$  is a time interval for the particle  $p$  (with respect to the frame of reference fixed by our choice of primitives). It may seem that it would have been simpler to take  $\mathcal{I}(p)$  as the interval of proper time of the particle  $p$ . However, this approach would complicate the treatment of systems of particles. In the main, the notion of proper time is most convenient in discussions restricted to the consideration of a single particle. From the remark in the previous section it is clear that it is not reasonable to require that the intervals  $\mathcal{I}(p)$  be overlapping. A second argument against such an assumption is the prominence in modern physics of elementary particles with very short life-spans.<sup>2/</sup> We note, however, that in studying certain special problems, such as that of defining a reasonable notion of center of mass of a S.R.P.M., it is desirable to restrict the discussion to systems in which  $\mathcal{I}(p) = (-\infty, +\infty)$  for every  $p$  in  $P$ .

If i) "c" is replaced by "1/k" in the inequality of Axiom A4 and the equation of Axiom A7, ii) k is treated as a primitive replacing c, and iii) Axiom A4 is modified to read: "The constant k is a non-negative real number such that...", then, by adding appropriate further axioms, we can get either classical or relativistic particle mechanics. Thus an additional axiom asserting that  $k=0$  gives us classical mechanics; and the assertion that  $k>0$  gives us relativistic mechanics.

We close this section with a number of definitions which will be useful later.

For  $p$  in  $P$  and  $t$  in  $\mathcal{A}(p)$ , we set

$$v_p(t) = (Ds_p)(t) \quad ;$$

$v_p(t)$  is, of course, the velocity of  $p$  at time  $t$ . With respect to a fixed element  $t_0$  in  $\mathcal{A}(p)$ , we define the function  $\tau_{t_0}$  (for  $p$  in  $P$  and  $t$  in  $\mathcal{A}(p)$ ) as follows:

$$\tau_{t_0} = \int_{t_0}^t \sqrt{1 - \frac{|v_p(t)|^2}{c^2}} dt \quad ;$$

$\tau_{t_0}(p, t)$  is the proper time of  $p$ . Since we are only interested in the derivative of this function with respect to  $t$ , and since the derivative is independent of  $t_0$ , we shall usually drop the subscript.

For  $p$  in  $P$  and  $t$  in  $\mathcal{A}(p)$ , we define the function  $q$  as follows:

$$q(p, t) = \langle s(p, t), t \rangle \quad .$$

It is natural to call  $q$  the space-time function.

For  $p$  in  $P$ ,  $t$  in  $\mathcal{A}(p)$  and  $i$  any positive integer, we define what we call the relativistic force function  $f^{\text{rel}}$  as follows:

$$f^{\text{rel}}(p, t, i) = \left\langle f(p, t, i), \frac{f(p, t, i) \cdot v_p(t)}{c^2} \right\rangle \quad .$$

Although it is not usual to adopt a special name for this function, the function itself is used frequently in textbook treatments of relativity.

By a c-particle path (for any positive number  $c$ ) we mean a set  $\mathcal{L}$  of points (i.e., vectors) in  $E_{n+1}$  for which there exists a S.R.P.M.  $\langle \{1\}, \mathcal{T}, m, s, f, c \rangle$  such that for every point  $X$  of  $E_{n+1}$ ,  $X$  is in  $\mathcal{L}$  if and only if there exists a  $t$  in  $\mathcal{T}(1)$  such that  $X = \langle s(1, t), t \rangle$ .<sup>3/</sup> It is obvious that if  $g$  is any twice-differentiable function defined on an interval  $T$  of real numbers and taking vectors in  $E_n$  as values, then the set of vectors  $\langle g(t), t \rangle$  for  $t$  in  $T$  is a c-particle path, provided that  $|(Dg)(t)| \leq c$  for all  $t$  in  $T$ .

By the slope of a line  $\alpha$  in  $E_{n+1}$ , whose projection on the  $(n+1)^{st}$ -axis is a non-degenerate segment, we mean the  $n$ -dimensional vector  $W$  such that for any two distinct points  $\langle z_1, x_1 \rangle$  and  $\langle z_2, x_2 \rangle$  of  $\alpha$

$$\frac{z_1 - z_2}{x_1 - x_2} = W.$$

By the speed of  $\alpha$  we mean the non-negative number  $|W|$ . By a c-inertial path we mean a line in  $E_{n+1}$  whose speed is less than  $c$ . We note that every segment of a c-inertial path is a c-particle path, but is not necessarily a c-inertial path (since a c-inertial path must be a whole line). By a c-line we mean a line in  $E_{n+1}$  whose speed is equal to  $c$ . The notion of a c-line corresponds to the intuitive notion of a light line.

If we want to refer to a S.R.P.M.  $\Gamma$  with numerical constant  $c$ , we shall write: S.R.P.M.  $\Gamma_c$ .

#### 4. Transformation Theorems.

We begin by defining the notion of a generalized Lorentz matrix. An intuitive discussion of such matrices follows Theorem 1.

Definition 1. Let  $c$ ,  $c'$  and  $\lambda$  be positive real numbers. Then a matrix  $A$  of order  $n+1$  is said to be a GENERALIZED LORENTZ MATRIX WITH RESPECT TO  $\langle c, c', \lambda \rangle$  if and only if there exist numbers  $\delta$  and  $\beta$ , an  $n$ -dimensional vector  $U$  and an orthogonal matrix  $\mathcal{E}$  of order  $n$  such that

$$\delta^2 = 1$$

$$\beta^2(1 - \frac{U^2}{c'^2}) = 1, \quad ,$$

and

$$(i) \quad A = \lambda \begin{pmatrix} 1 & 0 \\ 0 & \frac{c}{c'} \end{pmatrix} \begin{pmatrix} \mathcal{E} & 0 \\ 0 & \delta \end{pmatrix} \begin{pmatrix} 1 + \frac{\beta-1}{U^2} U^* U & -\frac{\beta U^*}{c'^2} \\ -\beta U & \beta \end{pmatrix}.$$

The following two lemmas simplify the statement and proof of Theorem 1.

Lemma 1. Let  $\langle \{1\}, \mathcal{T}, m, s, f, c \rangle$  be a S.R.P.M., let  $c'$  and  $\lambda$  be positive real numbers, and let  $A$  be a generalized Lorentz matrix with respect to  $\langle c, c', \lambda \rangle$ . Let the function  $h$  be defined by the equation (for every  $t$  in  $\mathcal{T}(1)$ )

$$h(t) = [ \langle s_1(t), t \rangle A ]_{n+1}.$$

Then the function  $Dh$  exists; its values are either always positive or always negative; and the function  $h$  is 1-1.

Proof. From Definition 1 and the hypothesis of the lemma we see that there are numbers  $\delta$  and  $\beta$ , an  $n$ -dimensional vector  $U$  and an orthogonal matrix  $\mathcal{E}$  such that

$$\delta^2 = 1$$

$$\beta^2(1 - \frac{U^2}{c'^2}) = 1$$

and

$$A = \begin{pmatrix} \lambda(\mathcal{E} + \frac{(\beta-1)\mathcal{E}U^*U}{U^2}) & -\frac{\lambda\beta\mathcal{E}U^*}{c'^2} \\ -\frac{\lambda c\delta\beta U}{c'} & \frac{\lambda c\delta\beta}{c'} \end{pmatrix}.$$

Thus,

$$h(t) = \frac{\lambda c \delta \beta t}{c'} - \frac{\lambda \beta s_1(t) \xi U^*}{c'^2} = \left( \frac{\lambda c \delta \beta}{c'} \right) (t - \frac{\delta s_1(t) \xi U^*}{cc'})$$

Hence

$$\frac{c'(Dh)(t)}{\delta \lambda \beta c} = 1 - \frac{v_1(t) \delta \xi U^*}{cc'} \geq 1 - \frac{|v_1(t) \xi| |U|}{cc'}$$

Using Axiom A4 and the fact that  $\xi$  is orthogonal, we have

$$\frac{c'(Dh)(t)}{\delta \lambda \beta c} \geq 1 - \frac{|U|}{c'}$$

Since  $|U| < c'$ , the function  $Dh$  is bounded away from zero, and it thus follows from Rolle's theorem that  $h$  is 1-1.

The following lemma is a theorem of matrix theory.

Lemma 2. Let  $c$ ,  $c'$  and  $\lambda$  be positive real numbers. Then a matrix  $a$  of order  $n+1$  is a generalized Lorentz matrix with respect to  $\langle c, c', \lambda \rangle$  if and only if

$$(i) \quad a \begin{pmatrix} 1 & 0 \\ 0 & -c'^2 \end{pmatrix} a^* = \lambda^2 \begin{pmatrix} 1 & 0 \\ 0 & -c^2 \end{pmatrix}$$

Proof. The proof of necessity is obtained by direct application of Definition 1.

For the proof of sufficiency, let

$$a = \begin{pmatrix} \mathcal{N} & K^* \\ L & m \end{pmatrix},$$

where  $\mathcal{N}$  is a matrix of order  $n$ ,  $K$  and  $L$  are  $n$ -dimensional vectors and  $m$  is a real number. From (i) we obtain at once:

$$(1) \quad \mathcal{N} \mathcal{N}^* - c'^2 K^* K = \lambda^2 1$$

$$(2) \quad \mathcal{N} L^* - c'^2 m K^* = 0$$

$$(3) \quad L L^* - c'^2 m = -\lambda^2 c^2$$

From (3) it follows that

$$(4) \quad m \neq 0$$

We define:

$$(5) \quad \beta = \frac{c' |m|}{c \lambda}$$

$$(6) \quad \delta = \frac{m}{|m|}$$

$$(7) \quad U = - \frac{L}{m}$$

$$(8) \quad \xi = \frac{1}{\lambda} \left( \mathcal{H} + \frac{(\beta-1)c'^2 K^* U}{\beta U^2} \right)$$

Since the right member of equation (i) of Definition 1 can be written

$$\begin{pmatrix} \lambda \left( \xi + \frac{(\beta-1)\xi U^* U}{U^2} \right) - \frac{\lambda \beta \xi U^*}{c'^2} \\ - \frac{\lambda c \delta \beta U}{c'} \quad \frac{\lambda c \delta \beta}{c'} \end{pmatrix},$$

in order to complete the proof it suffices to show that

$$(I) \quad \lambda \left( \xi + \frac{(\beta-1)\xi U^* U}{U^2} \right) = \mathcal{H}$$

$$(II) \quad - \frac{\lambda \beta \xi U^*}{c'^2} = K^*$$

$$(III) \quad - \frac{\lambda c \delta \beta U}{c'} = L$$

$$(IV) \quad \frac{\lambda c \delta \beta}{c'} = m$$

$$(V) \quad \delta^2 = 1$$

$$(VI) \quad \beta^2 \left( 1 - \frac{U^2}{c'^2} \right) = 1$$

$$(VII) \quad \xi \xi^* = 1$$

Equation (III) follows immediately from (5), (6), and (7), equation (IV) from (5) and (6), equation (V) from (6), and equation (VI) from (3), (5), and (7).

From (2) and (7) we get

$$(9) \quad \mathcal{H} U^* = -c'^2 K^*,$$

and then from (8) and (9) we have

$$\begin{aligned}
 (10) \quad \mathcal{E}\mathcal{E}^* &= \frac{1}{\lambda^2} \left( \mathcal{N} + \frac{(\beta-1)c'^2 K^* U}{\beta U^2} \right) \left( \mathcal{N}^* + \frac{(\beta-1)c'^2 U^* K}{\beta U^2} \right) \\
 &= \frac{1}{\lambda^2} \left\{ \mathcal{N}\mathcal{N}^* + \frac{(\beta-1)c'^2}{\beta^2 U^2} \left[ \beta (-c'^2 K^*) K + \beta K^* (-c'^2 K) \right. \right. \\
 &\quad \left. \left. + (\beta-1)c'^2 K^* K \right] \right\} \\
 &= \frac{1}{\lambda^2} \left[ \mathcal{N}\mathcal{N}^* + \frac{(\beta-1)c'^4}{\beta^2 U^2} K^* K (-2\beta + \beta - 1) \right] \\
 &= \frac{1}{\lambda^2} \left[ \mathcal{N}\mathcal{N}^* - \frac{(\beta^2-1)c'^4}{\beta^2 U^2} K^* K \right] .
 \end{aligned}$$

From (VI), (1) and (10) we conclude that

$$\mathcal{E}\mathcal{E}^* = \mathcal{I} ,$$

which establishes equation (VII). Multiplying both sides of (8) on the right by  $-\frac{\lambda\beta U^*}{c'^2}$ , and using (9), we get equation (II). Equation (I) follows from (8) and (II), which completes the proof of the lemma.

The following theorem is a generalization of the well-known result that the relativistic equation of motion is covariant under a Lorentz transformation.

Theorem 1. Let  $\langle P, \mathcal{T}, m, s, f, c \rangle$  be an  $n$ -dimensional S.R.P.M. Let  $c'$ ,  $\gamma$  and  $\lambda$  be positive real numbers, let  $B$  be an  $n+1$ -dimensional vector and let  $A$  be a generalized Lorentz matrix with respect to  $\langle c, c', \lambda \rangle$ . For each  $p$  in  $P$  let the function  $h_p$  be defined as follows (for all  $t$  in  $\mathcal{T}(p)$ ):

$$h_p(t) = [\langle s_p(t), t \rangle A + B]_{n+1} .$$

(By Lemma 1 the inverse function  $h_p^{-1}$  exists.) Let the function  $\mathcal{T}'$  be defined as follows: for  $p$  in  $P$ ,  $\mathcal{T}'(p)$  is the range of the function  $h_p$ ; and let the functions  $m'$ ,  $s'$  and  $f'$  be defined by the following equations (for  $p$  in  $P$ ,  $t'$  in  $\mathcal{T}'(p)$  and  $i$  in  $I$ ):

$$m'(p) = \gamma m(p) ,$$

$$s'(p, t') = [\langle s(p, h_p^{-1}(t')), h_p^{-1}(t') \rangle A + B]_{1, \dots, n} ,$$



$$f'(p, t', i) = \frac{\lambda c'^2}{\lambda^2 c^2} \left[ \left\langle f(p, h_p^{-1}(t'), i), \frac{f(p, h_p^{-1}(t'), i) \cdot v_p(h_p^{-1}(t'))}{c^2} \right\rangle a_{1, \dots, n} \right]$$

Then  $\Gamma' = \langle P, \mathcal{I}', m', s', f', c' \rangle$  is an n-dimensional S.R.P.M.

Proof. It will suffice to show that  $\Gamma'$  satisfies Axioms A4 and A7.

since the proof for the other axioms is trivial. Let

$$a = \begin{pmatrix} \mathcal{D} & E^* \\ F & g \end{pmatrix}$$

It is easy to show that for  $p$  in  $P$  and  $t'$  in  $\mathcal{I}'(p)$

$$(1) \quad v'_p(t') = \frac{\left\langle v_p(h_p^{-1}(t'), 1) \right\rangle \begin{pmatrix} \mathcal{D} \\ F \end{pmatrix}}{\left\langle v_p(h_p^{-1}(t'), 1) \right\rangle \begin{pmatrix} E^* \\ g \end{pmatrix}},$$

with the denominator of the right member of (1) always unequal to zero.

(Since in this proof we always consider a fixed particle  $p$ , we drop the subscript "p" from this point on.)

We have from Axiom A4

$$(2) \quad \lambda^2(|v(h^{-1}(t'))|^2 - c^2) < 0,$$

but

$$\lambda^2(|v(h^{-1}(t'))|^2 - c^2) = \lambda^2 \left\langle v(h^{-1}(t')), 1 \right\rangle \begin{pmatrix} 0 & 0 \\ 0 & -c^2 \end{pmatrix} \left\langle v(h^{-1}(t')), 1 \right\rangle^*.$$

Then by Lemma 2 we have

$$(3) \quad \lambda^2(|v(h^{-1}(t'))|^2 - c^2) = \left\langle v(h^{-1}(t')), 1 \right\rangle a \begin{pmatrix} 0 & 0 \\ 0 & -c'^2 \end{pmatrix} a^* \left\langle v(h^{-1}(t')), 1 \right\rangle^*.$$

The right member of (3) is equal to

$$(4) \quad \left\langle v(h^{-1}(t')), 1 \right\rangle \begin{pmatrix} \mathcal{D} \\ F \end{pmatrix} (\mathcal{D}^* F^*) \left\langle v(h^{-1}(t')), 1 \right\rangle^* - c'^2 \left\langle v(h^{-1}(t')), 1 \right\rangle \begin{pmatrix} E^* \\ g \end{pmatrix} \langle E_g \rangle \left\langle v(h^{-1}(t')), 1 \right\rangle^*,$$

and using (1) we see that (4) is equal to

$$(5) \quad (\left\langle v(h^{-1}(t')), 1 \right\rangle \begin{pmatrix} E^* \\ g \end{pmatrix} v'(t'))^2 - c'^2 (\left\langle v(h^{-1}(t')), 1 \right\rangle \begin{pmatrix} E^* \\ g \end{pmatrix})^2.$$

From (2), (3), (4), and (5) we conclude that

$$|v'(t')|^2 - c'^2 < 0 ,$$

which verifies Axiom A4 for  $\Gamma'$ .

It is not difficult to show that from Axiom A7 we have

$$(6) \quad \frac{m(p)}{\sqrt{1 - \frac{|v(t)|^2}{c^2}}} D\left[\frac{Dq}{\sqrt{1 - \frac{|v|^2}{c^2}}}\right](t) = \sum_{i=1}^{\infty} f^{rel}(p, t, i) .$$

Setting  $q'(t') = \langle s'(t'), t' \rangle$  for all  $t'$  in  $\mathcal{T}'(p)$ , we conclude from the hypothesis of our theorem that

$$q'(h(t)) = q(t)A + B ,$$

and thus

$$(7) \quad ((Dq') \circ h)(t)(Dh)(t) = (Dq)(t)A .$$

Directly from the definition of  $q$  and  $q'$  we obtain

$$(8) \quad (Dq)(t) \begin{pmatrix} 1 & 0 \\ 0 & -c^2 \end{pmatrix} ((Dq)(t))^* = |v(t)|^2 - c^2 ,$$

and

$$(9) \quad ((Dq') \circ h)(t) \begin{pmatrix} 1 & 0 \\ 0 & -c'^2 \end{pmatrix} (((Dq') \circ h)(t))^* = |(v' \circ h)(t)|^2 - c'^2 .$$

Using Lemma 1, Lemma 2, and (7) we obtain from (8) and (9)

$$|(v' \circ h)(t)|^2 - c'^2 = \frac{\lambda^2}{((Dh)(t))^2} (|v(t)|^2 - c^2) ;$$

and thus

$$(10) \quad 1 - \frac{|(v' \circ h)(t)|^2}{c'^2} = \frac{\lambda^2 c^2}{c'^2 ((Dh)(t))^2} \left(1 - \frac{|v(t)|^2}{c^2}\right) .$$

By Lemma 1,  $(Dh)(t)$  is either always positive or always negative; the remainder of our proof is analogous in the two cases, so that we shall only consider the case where it is always positive. We then have from (7) and (10)

$$\frac{[(Dq') \circ h](t)}{\sqrt{1 - \frac{|(v' \circ h)(t)|^2}{c'^2}}} = \frac{c' (Dq)(t) A}{\lambda c \sqrt{1 - \frac{|v(t)|^2}{c^2}}} ,$$

and hence

$$(11) \quad \left( \left[ \frac{Dq'}{\sqrt{1 - \frac{|v'|^2}{c'^2}}} \right] \circ h \right) (t) = \frac{c'(Dq)(t)Q}{\lambda c \sqrt{1 - \frac{|v(t)|^2}{c^2}}}.$$

Differentiating both sides of (11), and using (6), we obtain

$$(12) \quad (Dh)(t) \left( \left[ D \left( \frac{Dq'}{\sqrt{1 - \frac{|v'|^2}{c'^2}}} \right) \right] \circ h \right) (t) \\ = \frac{c'}{\lambda c} D \left( \frac{Dq}{\sqrt{1 - \frac{|v|^2}{c^2}}} \right) (t) Q \\ = \left( \sqrt{1 - \frac{|v(t)|^2}{c^2}} \right) \left( \frac{c'}{\lambda c m(p)} \right) \left( \sum_{i=1}^{\infty} f^{rel}(p, t, i) Q \right).$$

From (10), (12), and the hypothesis of our theorem we infer that

$$(13) \quad \gamma^{m(p)} \left( \left[ D \left( \frac{Dq'}{\sqrt{1 - \frac{|v'|^2}{c'^2}}} \right) \right] \circ h \right) (t) \\ = \sqrt{1 - \frac{|(v' \circ h)(t)|^2}{c'^2}} \sum_{i=1}^{\infty} f^{rel'}(p, h(t), i),$$

and from (13) we conclude immediately that Axiom A7 holds for  $\Gamma'$ .  $\square$

**Remark 1.** All the transformations mentioned in Definition 1 and Theorem 1 have a clear intuitive interpretation if we consider  $\langle P, \mathcal{T}, m, s, f, c \rangle$  as a physical system whose mechanical properties are observed and measured with respect to some (inertial) frame of reference and some set of units of measurement, and  $\langle P, \mathcal{T}', m', s', f', c' \rangle$  as the same physical system observed and measured with respect to some other (inertial) frame of reference and some other set of units of measurement. Thus,  $c$  is the old and  $c'$  the new velocity of light. The introduction of the number  $\gamma$  amounts to changing the unit of mass by an amount

$1/\gamma$ , and the vector  $B$  corresponds to shifting the origin of the spatial frame of reference by  $-[B]_1, \dots, n$ , and the origin of time by an amount  $-[B]_{n+1}$ . The number  $\lambda$  represents a uniform stretch of space and time. When  $\delta = -1$ , we have a reversal of the direction of time. The matrix  $\mathcal{C}$  represents (for  $n \leq 3$ ) a rotation of the spatial coordinates or a rotation followed by a reflection. The vector  $U$  represents the relative velocity of the two inertial frames of reference, and the number  $\beta$ , which is determined by  $U$  and  $c'$ , is the well-known Lorentz contraction factor. Finally, it is easy to check that the last matrix in the factorization of the matrix  $A$  yields the ordinary Lorentz transformations. We note that the rather complicated transformation of the forces is the velocity-dependent transformation to be expected in relativistic mechanics.

Remark 2. Theorem 2, our main theorem, is a sort of converse of Theorem 1: roughly speaking, we show that the transformations described in Theorem 1 are the only transformations which always take systems of relativistic particle mechanics into systems of relativistic particle mechanics. To facilitate the formulation and proof of Theorem 2, an additional lemma and some definitions will be useful.

Lemma 3. Let  $X_1 = \langle Z_1, x_1 \rangle$ ,  $X_2 = \langle Z_2, x_2 \rangle$  and  $X_3 = \langle Z_3, x_3 \rangle$  be any three points in  $E_{n+1}$  such that (i)  $x_1 < x_2 < x_3$ , (ii) there is a  $c$ -inertial path through  $X_1$  and  $X_2$ , and (iii) there is a  $c$ -inertial path through  $X_2$  and  $X_3$ . Then there is a  $c$ -particle path through  $X_1$ ,  $X_2$ , and  $X_3$ .

Proof. In view of the remark near the end of Section 3, it will suffice to construct a function  $g$  which: (a) is defined on the closed interval  $[x_1, x_3]$ ; (b) takes vectors in  $E_n$  as values; (c) is twice differentiable; (d) is such that for every  $t$  in  $[x_1, x_3]$ ,  $|(Dg)(t)| < c$ ; and (e) is such that

$$g(x_1) = Z_1 ,$$

$$g(x_2) = Z_2 ,$$

and

$$g(x_3) = Z_3 .$$

Let

$$a = x_2 - x_1$$

$$b = x_3 - x_2$$

$$V = \frac{Z_2 - Z_1}{a}$$

$$W = \frac{Z_3 - Z_2}{b}$$

$$\gamma = \frac{2c \log 2}{(c - \max(|V|, |W|)) \min(a, b)}$$

$$A = \frac{aV \log \cosh \gamma b + bW \log \cosh \gamma a}{a \log \cosh \gamma b + b \log \cosh \gamma a}$$

$$B = \frac{ab \gamma (W - V)}{a \log \cosh \gamma b + b \log \cosh \gamma a} .$$

The reader may verify that the function  $g$  defined by the following equation (for  $t$  in  $[x_1, x_3]$ ) has properties (a)-(e):

$$g(t) = Z_2 + (t - x_2)A + \frac{B}{\gamma} [\log \cosh \gamma (t - x_2)] .$$

Definition 2. Let  $\varphi_1$  be a function mapping  $R^+$  into  $R^+$ ; let  $\varphi_2$  be a function which is a 1-1 mapping of  $E_{n+1}$  into itself; and let  $\varphi_3$  be a function mapping  $E_{2n}$  into  $E_n$ . Then we call the ordered triple  $\langle \varphi_1, \varphi_2, \varphi_3 \rangle$  an ELIGIBLE TRANSFORMATION.

Definition 3. Let  $\bar{\Phi} = \langle \varphi_1, \varphi_2, \varphi_3 \rangle$  be an eligible transformation, let  $\Gamma = \langle P, \mathcal{T}, m, s, f, c \rangle$  be a S.R.P.M., and for each  $p$  in  $P$  let the function  $H_p$  be defined as follows (for every  $t$  in  $\mathcal{T}(p)$ )

$$H_p(t) = [\varphi_2(s(p, t), t)]_{n+1} .$$

Then by the  $\bar{\Phi}$ -TRANSFORM of  $\Gamma$  (which we also write:  $\bar{\Phi}(\Gamma)$ ), we mean the ordered quintuple  $\langle P, \mathcal{T}', m', s', f' \rangle$ , where for  $p$  in  $P$ :

$$m'(p) = \varphi_1(m(p)) \quad ;$$

$\mathcal{T}'(p)$  is the range of the function  $H_p$ ; and  $s'$  and  $f'$  are defined by the following equations for  $t'$  in  $\mathcal{T}'(p)$ , if the pre-image  $H_p^{-1}(t')$  of  $t'$  under  $H_p$  is unique, and otherwise they are undefined:

$$s'(p, t') = [\varphi_2(s(p, H_p^{-1}(t')), H_p^{-1}(t'))]_{1, \dots, n}$$

$$f'(p, t', i) = \varphi_3(f(p, H_p^{-1}(t'), i), v(p, H_p^{-1}(t'))) \quad ,$$

for  $i \geq 1$ .

We are now in a position to state and prove the main theorem of this paper.

Theorem 2.<sup>5/</sup> Let  $\Phi = \langle \varphi_1, \varphi_2, \varphi_3 \rangle$  be an eligible transformation, and let  $c$  and  $c'$  be positive real numbers such that (i) for every  $n$ -dimensional system of relativistic particle mechanics  $\Gamma_c$ ,  $\langle \Phi(\Gamma_c), c' \rangle$  is a system of relativistic particle mechanics, and (ii)  $\varphi_2$  carries no  $c$ -line into a  $c'$ -particle path. Then there exist positive real numbers  $\gamma$  and  $\lambda$ , an  $(n+1)$ -dimensional vector  $B$ , and a generalized Lorentz matrix  $A$  with respect to  $\langle c, c', \lambda \rangle$  such that for any vectors  $Z_1$  and  $Z_2$  in  $E_n$  with  $|Z_2| < c$ , every  $x$  in  $R$  and  $y$  in  $R^+$ ,

$$\varphi_1(y) = \gamma y \quad ,$$

$$\varphi_2(Z_1, x) = \langle Z_1, x \rangle A + B \quad ,$$

$$\varphi_3(Z_1, Z_2) = \frac{\gamma c'^2}{\lambda^2 c^2} [\langle Z_1, \frac{Z_1 \cdot Z_2}{c^2} \rangle A]_{1, \dots, n} \quad .$$

Proof. We first want to show that if  $Z$  is any vector in  $E_n$  such that  $|Z| < c$ , then

$$\varphi_3(0, Z) = 0 \quad .$$

Setting  $P = \{1\}$ ,  $\mathcal{T}(1) = (-\infty, \infty)$ ,  $m(1) = 1$ , and for  $t$  in  $\mathcal{T}(1)$

$$s(1, t) = Zt \quad ,$$

$$f(1, t, i) = 0 \quad \text{for } i \geq 1 \quad ,$$

we see that  $\langle P, \mathcal{T}, m, s, f, c \rangle$  is a S.R.P.M. Since for every  $t$  in  $\mathcal{T}(1)$ ,

$Z = v(1, t)$ , we conclude from the hypothesis of our theorem, Definition 2, and

Axiom A6 that the series

$$\varphi_3(0, Z) + \varphi_3(0, Z) + \dots$$

is absolutely convergent. Hence,

$$(1) \quad \varphi_3(0, Z) = 0.$$

For every segment  $\Delta$  of a c-inertial path there exists a one-particle S.R.P.M.  $\langle \{1\}, \mathcal{A}_{m,s,f,c} \rangle$  such that for every  $t$  in  $\mathcal{A}(1)$

$$f(p, t, i) = 0 \quad \text{for } i \geq 1,$$

and for every vector  $X$  in  $E_{n+1}$ ,  $X$  is in  $\Delta$  if and only if there is a  $t$  in  $\mathcal{A}(1)$  such that

$$X = \langle s(1, t), t \rangle.$$

Hence, it follows immediately from (1) and the hypothesis of our theorem that

(2)  $\varphi_2$  carries segments of c-inertial paths into segments of c'-inertial paths.

Let  $\Gamma_c = \langle P, \mathcal{A}_{m,s,f,c} \rangle$  be any S.R.P.M. with constant  $c$ . By hypothesis  $\langle \bar{\Phi}(\Gamma_c), c \rangle$  is a S.R.P.M. For any  $p$  in  $P$ , if  $t_1$  and  $t_2$  are in  $\mathcal{A}(p)$  and  $t_1 \neq t_2$ , then

$$\varphi_2(s(p, t_1), t_1) \neq \varphi_2(s(p, t_2), t_2),$$

since  $\varphi_2$  is 1-1. Suppose now that

$$[\varphi_2(s(p, t_1), t_1)]_{n+1} = [\varphi_2(s(p, t_2), t_2)]_{n+1}.$$

Then we must have

$$[\varphi_2(s(p, t_1), t_1)]_{1, \dots, n} \neq [\varphi_2(s(p, t_2), t_2)]_{1, \dots, n};$$

but then  $\langle \bar{\Phi}(\Gamma_c), c' \rangle$  is not a S.R.P.M., for  $p$  is required to be in two places at the same time, which violates Axiom A3. We thus conclude:

(3)  $\varphi_2$  is 1-1 in the last coordinate along the space-time path of any particle of a S.R.P.M.  $\Gamma_c$ , and thus the pre-image under  $\varphi_2$  of any point  $t'$  in  $\mathcal{A}'(p)$  is unique.

Furthermore, since by hypothesis  $\varphi_2$  takes the interval  $\mathcal{A}(p)$  into an interval  $\mathcal{A}'(p)$ , we have:

- (4)  $\varphi_2$  is continuous in the last coordinate along the space-time path of any particle of a S.R.P.M.

From (4) and the fact that any two points  $\langle Z, x \rangle$  and  $\langle Z, y \rangle$  lie on a c-inertial path, we obtain:

- (5) For any point  $\langle Z, x \rangle$  and every  $\epsilon > 0$ , there exists a  $\delta > 0$  such that for any point  $\langle Z, y \rangle$  if  $|x - y| < \delta$ , then  $|x' - y'| < \epsilon$ , where  $x' = [\varphi_2(Z, x)]_{n+1}$  and  $y' = [\varphi_2(Z, y)]_{n+1}$ .

We next show that

- (I)  $\varphi_2$  is continuous.

Let  $\langle Z_1, x_1 \rangle$  be any point of  $E_{n+1}$ , and let  $\epsilon$  be any positive number. Let  $\epsilon^* = \frac{\epsilon}{2(1+c)}$ . Using (5), let  $\delta^*$  be a positive number such that if  $|x_1 - y| < \delta^*$  then  $|x'_1 - y'| < \epsilon^*$ , where  $x'_1 = [\varphi_2(Z_1, x_1)]_{n+1}$  and  $y' = [\varphi_2(Z_1, y)]_{n+1}$ ; and let  $\delta = \frac{c\delta^*}{3c+2}$ . We shall show that if  $\langle Z_2, x_2 \rangle$  is any point of  $E_{n+1}$  such that

$$(6) \quad |\langle Z_1, x_1 \rangle - \langle Z_2, x_2 \rangle| < \delta,$$

then  $|\varphi_2(Z_1, x_1) - \varphi_2(Z_2, x_2)| < \epsilon$ . Suppose for definiteness that

$$(7) \quad x_1 \geq x_2.$$

We may choose  $x_0$  and  $x_3$  so that

$$(8) \quad x_2 - \frac{|Z_2 - Z_1|}{c} - \delta < x_0 < x_2 - \frac{|Z_2 - Z_1|}{c},$$

and

$$(9) \quad x_1 + \frac{|Z_2 - Z_1|}{c} < x_3 < x_1 + \frac{|Z_2 - Z_1|}{c} + \delta.$$

From (7), (8), and (9), we obtain

$$(10) \quad |x_3 - x_0| < |x_2 - x_1| + 2 \frac{|Z_2 - Z_1|}{c} + 2\delta;$$

and from (6) and (10) we then infer that

$$(11) \quad |x_3 - x_0| < \frac{(3c+2)\delta}{c} = \delta^*.$$

Since from (7), (8), and (9) we have

$$(12) \quad x_0 < x_2 \leq x_1 < x_3,$$



we obtain from (11)

$$\begin{aligned} |x_3 - x_1| &< \delta^* \\ |x_1 - x_0| &< \delta^* \end{aligned} .$$

Consequently, by (5)

$$|x'_3 - x'_1| + |x'_1 - x'_0| < 2\epsilon^* ,$$

and thus by the triangle inequality

$$(13) \quad |x'_3 - x'_0| < 2\epsilon^* ,$$

where

$$\begin{aligned} x'_0 &= [\varphi_2(z_1, x_0)]_{n+1} , \\ x'_3 &= [\varphi_2(z_1, x_3)]_{n+1} . \end{aligned}$$

From the second part of (8) it follows that there is a c-inertial path through  $\langle z_1, x_0 \rangle$  and  $\langle z_2, x_2 \rangle$ ; and from (7) and the first part of (9) it follows that there is a c-inertial path through  $\langle z_2, x_2 \rangle$  and  $\langle z_1, x_3 \rangle$ . We thus conclude from Lemma 3 that there exists a c-particle path through  $\langle z_1, x_0 \rangle$ ,  $\langle z_2, x_2 \rangle$  and  $\langle z_1, x_3 \rangle$ . As before, for abbreviation, we set

$$\begin{aligned} x'_2 &= [\varphi_2(z_2, x_2)]_{n+1} , \\ z'_2 &= [\varphi_2(z_2, x_2)]_{1, \dots, n} , \\ z'_{1i} &= [\varphi_2(z_1, x_i)]_{1, \dots, n} \quad \text{for } i=0,1,3. \end{aligned}$$

Since  $\varphi_2$  is 1-1 and continuous in the last coordinate along any c-particle path, it is monotone in the last coordinate along any c-particle path, and we thus have: either

$$(14) \quad \left\{ \begin{array}{ll} \begin{array}{l} x'_0 < x'_1 < x'_3 \\ x'_0 < x'_2 < x'_3 \end{array} & , \\ \text{or} & ; \\ \begin{array}{l} x'_3 < x'_1 < x'_0 \\ x'_3 < x'_2 < x'_0 \end{array} & . \end{array} \right.$$

Also, since segments of  $c$ -inertial paths are carried by  $\varphi_2$  into segments of  $c'$ -inertial paths, we have:

$$(15) \quad |z'_{11} - z'_2| \leq |z'_{11} - z'_{13}| + |z'_{13} - z'_2| < c'|x'_1 - x'_3| + c'|x'_3 - x'_2| ,$$

and

$$(16) \quad |z'_{11} - z'_2| \leq |z'_{11} - z'_{10}| + |z'_{10} - z'_2| < c'|x'_1 - x'_0| + c'|x'_0 - x'_2| .$$

We obtain from (14), (15), and (16):

$$(17) \quad 2|z'_{11} - z'_2| < c'(|x'_3 - x'_1| + |x'_1 - x'_0| + |x'_3 - x'_2| + |x'_0 - x'_2|) \\ < 2c'|x'_3 - x'_0| .$$

Thus from (13) and (17) we conclude that

$$|z'_{11} - z'_2| < 2c'\epsilon^* ,$$

and from (13) and (14) that

$$|x'_1 - x'_2| < 2\epsilon^* ,$$

and since  $\epsilon^* = \frac{\epsilon}{2(1+c')}$ , we infer that

$$|\varphi_2(z_1, x_1) - \varphi_2(z_2, x_2)| < \epsilon ,$$

which establishes (I).

We next establish:

(II)  $\varphi_2$  carries parallel segments of  $c$ -inertial paths into parallel segments of  $c$ -inertial paths.

It is clearly sufficient to show that  $\varphi_2$  carries parallel  $c$ -inertial paths into parallel segments of  $c'$ -inertial paths. Let  $\eta_1$  and  $\eta_2$  be two parallel  $c$ -inertial paths, and let  $\eta_3$  be a  $c$ -inertial path which intersects  $\eta_1$  and  $\eta_2$  in the points  $A_1$  and  $A_2$  respectively (obviously such a  $c$ -inertial path  $\eta_3$  exists). (See Figure 1.) As previously, we use a prime to designate the image under  $\varphi_2$  of a point, line, etc. We may construct a fourth  $c$ -inertial path which intersects  $\eta_3$  between  $A_1$  and  $A_2$  and which intersects  $\eta_1$  and  $\eta_2$  at points distinct from  $A_1$  and  $A_2$ . Consequently, we infer from (2) that the

segments  $\eta'_1$  and  $\eta'_2$  lie in the same plane in the image space of  $\varphi_2$ . (See Figure 2.) Suppose now that  $\eta'_1$  and  $\eta'_2$  are not parallel. We extend (if necessary)  $\eta'_1$  and  $\eta'_2$  to their point of intersection, say,  $J'$ . We next select  $B'$  on  $\eta'_1$  between  $J'$  and  $A'_1$  (we use "between" in such a way that  $B'$  must be distinct from  $J'$  and  $A'_1$ ); similarly, we select  $D'$  on  $\eta'_2$  between  $J'$  and  $A'_2$ . We now consider the pre-images,  $B$  and  $D$ , of  $B'$  and  $D'$ . Since  $\varphi_2$  is 1-1 and continuous, it is clear that  $B$  and  $D$  must be on the same side of  $\eta_3$ , i.e., the segment  $BD$  does not intersect  $\eta_3$ . Let  $E$  be a point on  $\eta_3$  between  $A_1$  and  $A_2$ . Then, since  $\eta_3$  is a c-inertial path, one of the numbers  $[A_2]_{n+1} - [E]_{n+1}$  and  $[A_1]_{n+1} - [E]_{n+1}$  is positive, and the other is negative. Since  $\eta_1$  and  $\eta_2$  are parallel,  $[B]_{n+1} - [A]_{n+1}$  and  $[D]_{n+1} - [A_2]_{n+1}$  have the same sign. We then construct a line through  $D'$  parallel to  $\eta'_1$  or through  $B'$  parallel to  $\eta'_2$ , according to whether  $[A_2]_{n+1} - [E]_{n+1}$  or  $[A_1]_{n+1} - [E]_{n+1}$  agrees in sign with  $[B]_{n+1} - [A]_{n+1}$ . Suppose, for definiteness, (see Figure 1) that  $[A_2]_{n+1} - [E]_{n+1}$  agrees in sign and that this sign is positive. Let  $F'$  be the point of intersection of  $\eta'_3$  with the line through  $B'$  parallel to  $\eta'_1$ . By construction  $F'$  is between  $A'_1$  and  $A'_2$ , and thus  $F$  is between  $A_1$  and  $A_2$ .

We then have:

$$\begin{aligned} |[D]_{1,\dots,n} - [F]_{1,\dots,n}| &\leq |[D]_{1,\dots,n} - [A_2]_{1,\dots,n}| \\ &\quad + |[A_2]_{1,\dots,n} - [F]_{1,\dots,n}| \\ &< c([D]_{n+1} - [A_2]_{n+1}) + c([A_2]_{n+1} - [F]_{n+1}) \\ &< c([D]_{n+1} - [F]_{n+1}). \end{aligned}$$

Hence, the line through  $D$  and  $F$  is a c-inertial path. This line intersects  $\eta_1$  at a point, say,  $G$ , and, furthermore, by construction  $DFG$  is a segment of a c-inertial path, and hence the image  $D'F'G'$  is a segment of a c'-inertial path. But  $D'F'$  is parallel to  $\eta'_1$ , and the image of  $G$  does not lie on the

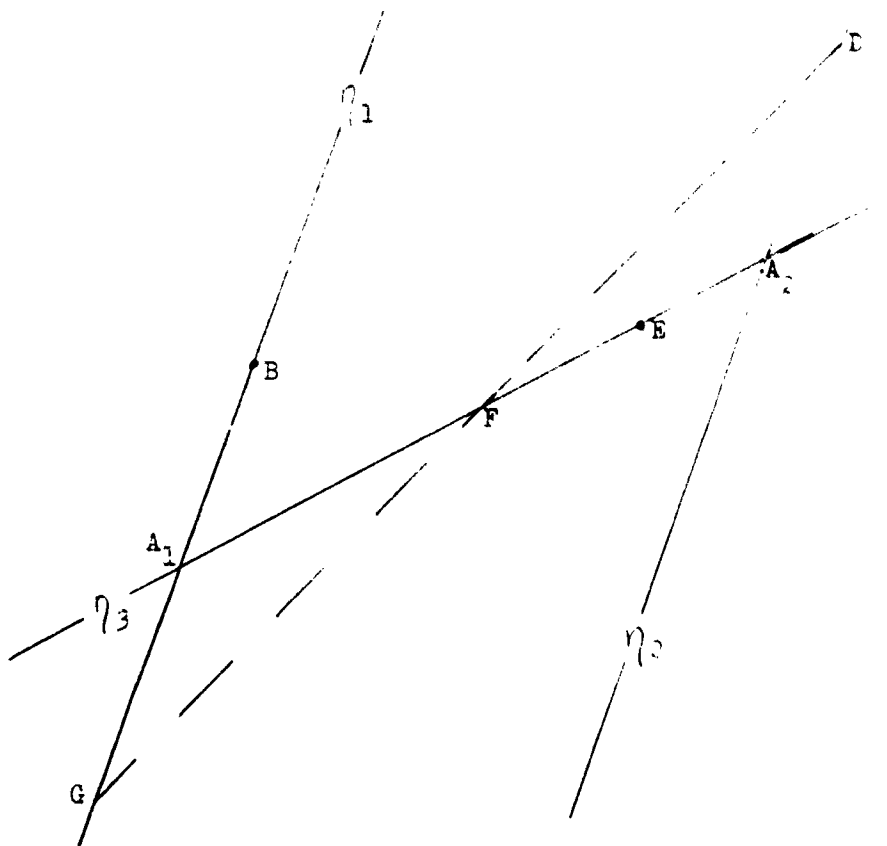


Figure 1

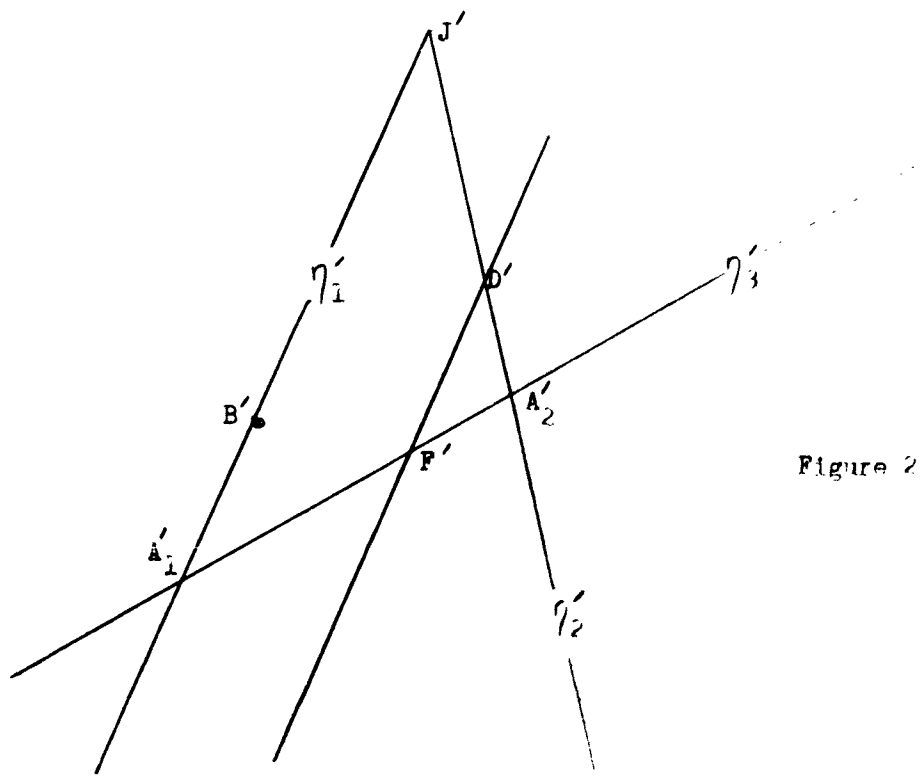


Figure 2

extension of  $D'F'$ , which is a contradiction. Thus  $\eta_1$  and  $\eta_2$  are parallel, and the proof of (II) is complete.

We next show that

(III)  $\Phi_2$  carries the midpoint of any finite segment  $\alpha$  of a c-inertial path into the midpoint of  $\alpha'$ .

We consider a fixed plane containing  $\alpha$  and a line parallel to the t-axis (the  $(n+1)^{st}$  coordinate axis). In this plane we construct, with  $\alpha$  as a diagonal, a parallelogram whose sides and other diagonal are segments of c-inertial paths. Let the speed of the c-inertial path containing  $\alpha$  be  $k$ . It is clear that through any point of our fixed plane there are exactly two lines with speed  $\ell$ , for every positive number  $\ell$ . Obviously, we may construct a parallelogram  $P$  with  $\alpha$  as one diagonal, with the other diagonal a segment of a c-inertial path with speed  $\frac{1}{4}(3k+c)$ , and with one side a segment of a c-inertial path with speed  $\frac{1}{2}(k+c)$ . The other side of the parallelogram  $P$  is then a segment of a c-inertial path with speed  $\frac{1}{2}(5k+c)$ . We conclude from (II) that  $P$  is carried by  $\Phi_2$  into a parallelogram  $P'$ , and the diagonals of  $P$  are carried into the diagonals of  $P'$ . Hence the midpoint of  $\alpha$  is carried into the midpoint of  $\alpha'$  and (III) is established.

We next show that

(IV)  $\Phi_2$  carries arbitrary lines into lines.

Let  $\alpha$  be an arbitrary line in  $E_{n+1}$  and let  $\langle z_1, x_1 \rangle$  and  $\langle z_2, x_2 \rangle$  be any two points on  $\alpha$ . We now construct an "inertial" parallelogram through these two points. For definiteness, we assume:

$$x_1 \geq x_2.$$

We set

$$z_0 = \frac{z_1 + z_2}{2},$$

and we choose  $x_0$  and  $x_3$  so that:

$$x_0 < x_2 - \frac{|z_1 - z_2|}{2c},$$

$$x_3 > x_1 + \frac{|z_1 - z_2|}{2c},$$

$$|\langle z_1, x_1 \rangle - \langle z_0, x_3 \rangle| = |\langle z_0, x_0 \rangle - \langle z_2, x_2 \rangle|$$

$$|\langle z_1, x_1 \rangle - \langle z_0, x_0 \rangle| = |\langle z_0, x_3 \rangle - \langle z_2, x_2 \rangle|$$

Let (see Figure 3)

$$A = \langle z_1, x_1 \rangle$$

$$B = \langle z_2, x_2 \rangle$$

$$C = \langle z_0, x_0 \rangle$$

$$D = \langle z_0, x_3 \rangle$$

$$E = \frac{A+B}{2}$$

$$F = \frac{B+D}{2}$$

$$G = \frac{A+C}{2}$$

$$H = \frac{A+D}{2}$$

$$K = \frac{B+C}{2}$$

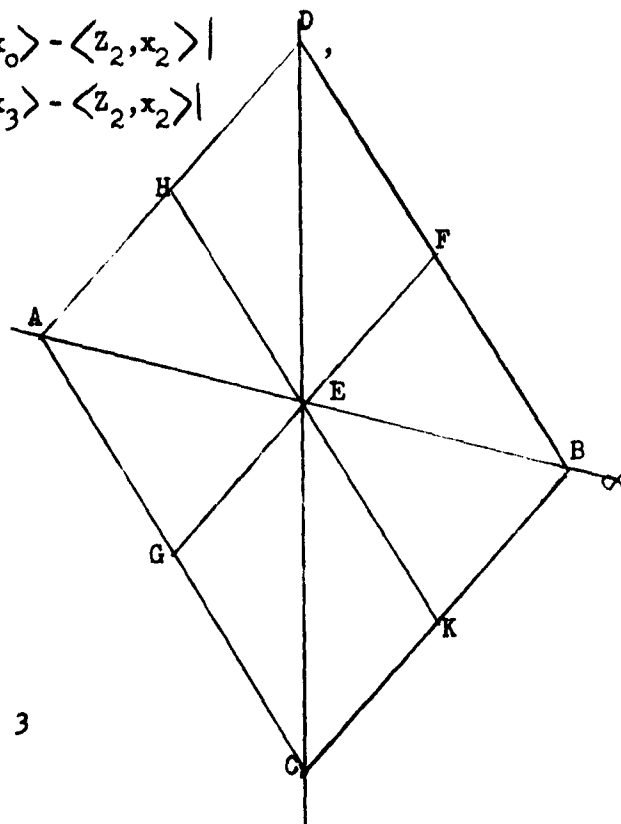


Figure 3

Since the sides of the parallelogram ACBD are by construction segments of  $c$ -inertial paths, we conclude from (II) that  $A'C'B'D'$  is a parallelogram, where  $A' = \Phi_2(A)$ , etc., and that the sides of  $A'C'B'D'$  are segments of  $c'$ -inertial paths. Moreover, it is clear that by construction CED, FEG and HEK are segments of  $c$ -inertial paths, and consequently  $C'E'D'$ ,  $F'E'G'$ , and  $H'E'K'$  are segments of  $c'$ -inertial paths. Hence by (III),  $F'$ ,  $G'$ ,  $H'$ , and  $K'$  are the midpoints of the respective sides of  $A'C'B'D'$ . Thus,  $E'$ , the point of intersection of the segments  $F'G'$  and  $H'K'$ , is the point of intersection of the diagonals of  $A'C'B'D'$ . Consequently,  $E'$  is the midpoint of the segment  $A'B'$ . Since midpoints of finite segments are carried into midpoints of finite segments and since  $\Phi_2$  is continuous, the proof of (IV) is complete.

From (IV) and the fact that  $\varphi_2$  is 1-1 and continuous, we immediately infer that  $\varphi_2$  is a projective transformation, and since it takes no finite point into a point at infinity, we conclude that

(18)  $\varphi_2$  is a non-singular affine transformation, that is, for every point  $\langle Z, x \rangle$  in  $E_{n+1}$

$$(19) \quad \varphi_2(Z, x) = \langle Z, x \rangle \cdot A + B,$$

where  $A$  is a non-singular matrix of order  $n+1$  and  $B$  is an  $(n+1)$ -dimensional vector.

Now let

$$(20) \quad A = \begin{pmatrix} D & E^* \\ F & g \end{pmatrix} \quad \text{and} \quad B = \langle B_1, b \rangle,$$

where  $D$  is a matrix of order  $n$ ,  $B_1$ ,  $E$  and  $F$  are  $n$ -dimensional vectors, and  $b$  and  $g$  are real numbers. Then

$$(21) \quad \varphi_2(Z, x) = \langle ZD + xF + B_1, ZE^* + gx + b \rangle.$$

Let  $\alpha$  be a c-line such that for any two distinct points  $\langle Z_1, x_1 \rangle$  and  $\langle Z_2, x_2 \rangle$  of  $\alpha$

$$\frac{Z_1 - Z_2}{x_1 - x_2} = W.$$

Obviously,  $|W| = c$ .  $\alpha$  is carried by  $\varphi_2$  into a line  $\alpha'$ . We want to show that  $\alpha$  is carried into a  $c'$ -line. From (21) it follows that the slope  $W'$  of  $\alpha'$  is given by

$$(22) \quad W' = \frac{W D + F}{W E^* + g}.$$

By the hypothesis of our theorem

$$(23) \quad |W'| \geq c'.$$

Consider now a sequence of c-inertial lines  $\alpha_1, \alpha_2, \dots$ , whose slopes  $W_1, W_2, \dots$  are such that

$$\lim_{i \rightarrow \infty} W_i = W.$$

From (21) and the hypothesis of our theorem we have

$$|W'_1| = \left| \frac{W_1 \partial + F}{W_1 E^* + g} \right| < c'.$$

Hence, if  $WE^* + g \neq 0$ , then

$$(24) \quad |W'| = \left| \frac{W \partial + F}{WE^* + g} \right| = \left| \lim_{i \rightarrow \infty} \frac{W_1 \partial + F}{W_1 E^* + g} \right| \leq c'.$$

Suppose now that  $WE^* + g = 0$ . Then

$$\lim_{i \rightarrow \infty} (W_1 E^* + g) = 0,$$

and therefore

$$\lim_{i \rightarrow \infty} (W_1 \partial + F) = 0.$$

Hence,

$$W \partial + F = 0,$$

but then

$$\langle W, 1 \rangle A = 0,$$

which is impossible, since  $A$  is non-singular. Thus we have

$$(25) \quad |W'| = c'.$$

For subsequent use we observe that for any S.R.P.M.,  $\langle P, \mathcal{T}_{m,s,f,c} \rangle$

and any  $p$  in  $P$  and  $t$  in  $\mathcal{T}(p)$

$$(26) \quad v(p,t)E^* + g \neq 0.$$

For  $v(p,t) \neq 0$ , the argument is the same as above; in case  $v(p,t) = 0$  for some  $t$ , on the supposition that  $v(p,t)E^* + g = 0$ , we must have  $g = 0$  and  $F = 0$ , which again contradicts the non-singularity of  $A$ .

From (22) and (25) we get

$$\frac{W \partial \partial^* W^* + 2W \partial F^* + |F|^2}{(WE^* + g)^2} = c'^2,$$

and hence

$$(27) \quad W(\partial \partial^* - c'^2 E^* E)W^* + 2W(\partial F^* - c'^2 E^* g) + |F|^2 - c'^2 g^2 = 0.$$

Since (27) holds for an arbitrary  $c$ -line, we may replace  $W$  by  $-W$ , and thus conclude that



$$W(\partial F^* - c'^2 E^* g) = 0 ;$$

and therefore, since the direction of  $W$  is arbitrary,

$$(28) \quad \partial F^* = c'^2 E^* g .$$

In view of the fact that (26) holds for  $v(p,t)=0$ , we have

$$g \neq 0 ,$$

and we may then obtain from (28)

$$(29) \quad E^* = \frac{\partial F^*}{c'^2 g} .$$

Using (20) and (29) we obtain

$$|A| = g \left| \partial \left( \partial - \frac{F^* F}{c'^2 g^2} \right) \right| ,$$

and since  $A$  is non-singular, we have

$$(30) \quad |F|^2 - c'^2 g^2 \neq 0 ;$$

and from (27), (28), and (30) it follows that

$$\partial \partial^* - c'^2 E^* E \neq 0 .$$

Thus, from (27), we have

$$W(\partial \partial^* - c'^2 E^* E)W^* = \left( \frac{c'^2 g^2 - |F|^2}{c^2} \right) |W|^2 .$$

Using again the fact that the direction of  $W$  is arbitrary, we infer that

$$(31) \quad \partial \partial^* - c'^2 E^* E = \frac{c'^2 g^2 - |F|^2}{c^2} I = \mu I .$$

where  $\mu = \frac{c'^2 g^2 - |F|^2}{c^2}$ . From (28) and (31) we obtain

$$(32) \quad A \begin{pmatrix} \partial & 0 \\ 0 & -c'^2 \end{pmatrix} A^* = \begin{pmatrix} \partial \partial^* - c'^2 E^* E & \partial F^* - c'^2 E^* g \\ (\partial F^* - c'^2 E^* g)^* & F F^* - c'^2 g^2 \end{pmatrix} \\ = \begin{pmatrix} \mu I & 0 \\ 0 & -\mu c^2 \end{pmatrix} = \mu \begin{pmatrix} I & 0 \\ 0 & -c^2 \end{pmatrix} .$$

We next want to show that  $\mu$  is positive. Let  $\langle z_1, x_1 \rangle$  and  $\langle z_2, x_2 \rangle$  be two points in  $E_{n+1}$  such that

$$\frac{|z_1 - z_2|}{|x_1 - x_2|} < c,$$

and let

$$V = \langle z_1, x_1 \rangle - \langle z_2, x_2 \rangle,$$

and

$$V' = VA.$$

From (32) we obtain

$$VA \begin{pmatrix} 1 & 0 \\ 0 & -c'^2 \end{pmatrix} A^* V^* = \mu V \begin{pmatrix} 1 & 0 \\ 0 & -c^2 \end{pmatrix} V^*.$$

Hence,

$$\mu = \frac{V' \begin{pmatrix} 1 & 0 \\ 0 & -c'^2 \end{pmatrix} (V')^*}{V \begin{pmatrix} 1 & 0 \\ 0 & -c^2 \end{pmatrix} V^*}.$$

By the hypothesis on  $V$ ,

$$V \begin{pmatrix} 1 & 0 \\ 0 & -c^2 \end{pmatrix} V^* < 0,$$

and from the fact that  $c$ -inertial paths are carried into  $c'$ -inertial paths, we have

$$V' \begin{pmatrix} 1 & 0 \\ 0 & -c'^2 \end{pmatrix} (V')^* < 0.$$

Thus,  $\mu$  is positive since it is the ratio of two negative numbers. We set

$$(33) \quad \lambda = \sqrt{\mu}.$$

We then conclude from (32), (33), Definition 1, and Lemma 2 that

$$(34) \quad A \text{ is a generalized Lorentz matrix with respect to } \langle c, c', \lambda \rangle.$$

We now turn to the function  $\Phi_3$  which transforms the forces. In deducing the form of  $\Phi_3$  it will be convenient to make use of the functions  $\tau$ ,  $q$  and  $f^{\text{rel}}$  defined in Section 3 (in the course of the present proof we obtain their transformation properties). It is also useful to introduce the function  $H$  defined by the following equation for every  $p$  in  $P$  and  $t$  in  $\mathcal{I}(p)$ ,

$$H(p, t) = [\varphi_2(s(p, t), t)]_{n+1}.$$

We thus have that for  $t'$  the element in  $\mathcal{A}'(p)$  corresponding to  $t$  in  $\mathcal{A}(p)$

$$H_p(t) = t'.$$

We obtain from (21)

$$(35) \quad D(H_p)(t) = v_p(t)E^* + g.$$

For any S.R.P.M.  $\Gamma = \langle P, \mathcal{A}, m, s, f, c \rangle$  the following equation is a direct consequence of Axiom 7 and the appropriate definitions (for any  $p$  in  $P$  and  $t$  in  $\mathcal{A}(p)$ )

$$(36) \quad m(p) \frac{d^2 q_p}{d\tau_p^2}(t) = \sum_{i=1}^{\infty} f^{\text{rel}}(p, t, i);$$

and also, under the hypothesis of our theorem,

$$(37) \quad \Phi_1(m(p)) \frac{d^2 q'_p}{d\tau_p'^2}(H_p(t)) = \sum_{i=1}^{\infty} f^{\text{rel}'}(p, H_p(t), i).$$

We now obtain the relationship between  $\frac{d^2 q'_p}{d\tau_p'^2}(H_p(t))$  and  $\frac{d^2 q_p}{d\tau_p^2}(t)$ .

Using (35), we obtain

$$(38) \quad \begin{aligned} \frac{d(\tau'_p \circ H_p)}{d\tau_p}(t) &= \left( \frac{D(\tau'_p \circ H_p)}{D\tau_p} \right)(t) \\ &= \frac{[(D\tau'_p)(H_p(t))][(DH_p)(t)]}{(D\tau_p)(t)} \\ &= \frac{\sqrt{1 - \frac{|v'_p(H_p(t))|^2}{c^2}} (v_p(t)E^* + g)}{\sqrt{1 - \frac{|v_p(t)|^2}{c^2}}}. \end{aligned}$$

It is easy to show that

$$(39) \quad v'_p(H_p(t)) = \frac{v_p(t)E^* + F}{v_p(t)E^* + g},$$

and hence, using (39) and squaring (38), we get

$$(40) \quad \left[ \left( \frac{D(\tau'_p \circ H_p)}{D\tau_p} \right) (t) \right]^2 = \frac{c^2}{c'^2} \left[ \frac{c'^2 (v_p(t)E^* + g)^2 - (v_p(t)D + F)^2}{c^2 - |v_p(t)|^2} \right]$$

Using (27) to give us the expansion of the right member of (40) and then using (32) to simplify the result, we obtain

$$(41) \quad \left[ \frac{D(\tau'_p \circ H_p)}{D\tau_p} (t) \right]^2 = \frac{c^2}{c'^2} \left[ \frac{-v_p(t) \lambda^2 \mathcal{Q}(v_p(t))^* + \lambda^2 c^2}{c^2 - |v_p(t)|^2} \right] \\ = \frac{c^2}{c'^2} \lambda^2 ;$$

and hence,

$$(42) \quad D(\tau'_p \circ H_p)(t) = \frac{\delta c \lambda}{c'} (D\tau_p)(t) ,$$

where  $\delta^2 = 1$ . We have from (21) and Definition 3

$$q'_p(H_p(t)) = q_p(t)A + B ,$$

and thus

$$(43) \quad [D(q'_p \circ H_p)(t)] = (Dq_p)(t)A$$

Since

$$\left[ \frac{D(q'_p \circ H_p)}{D(\tau'_p \circ H_p)} \right] (t) = \frac{(Dq'_p)(H_p(t))(DH_p)(t)}{(D\tau'_p)(H_p(t))(DH_p)(t)} \\ = \left( \frac{Dq'_p}{D\tau'_p} \right) (H_p(t)) ,$$

it is easily shown that

$$(44) \quad \frac{d^2 q'_p}{d\tau'^2_p} (H_p(t)) = \frac{D \left[ \frac{D(q'_p \circ H_p)}{D(\tau'_p \circ H_p)} \right]}{D(\tau'_p \circ H_p)} (t)$$

From (42), (43), and (44) we infer that

$$(45) \quad \frac{d^2 q'_p}{d\tau'^2_p} (H_p(t)) = \frac{D \left[ \frac{(Dq_p)A}{\frac{\delta c \lambda}{c'} (D\tau_p)} \right]}{\frac{\delta c \lambda}{c'} (D\tau_p)} (t) \\ = \frac{c'^2}{c^2 \lambda^2} \frac{d^2 q_p}{d\tau^2_p} (t) A$$

Now let  $X$  and  $Y$  be any two vectors in  $E_n$  with  $X \neq 0$  and  $|Y| < c$ . Then we set:

$$P = \{1\}$$

$$m(1) = 1$$

$$Z = (1 - \frac{Y^2}{c^2})(X - \frac{(X \cdot Y)Y}{c^2})$$

$$\mathcal{U}(1) = (-\frac{c-|Y|}{|Z|}, \frac{c-|Y|}{|Z|})$$

and for all  $t$  in  $\mathcal{U}(1)$ ,

$$s_1(t) = tY + \frac{1}{2} t^2 Z$$

$$f(1, t, 1) = \frac{Z}{1 - \frac{|v_1(t)|^2}{c^2}} + \frac{v_1(t)(Z \cdot v_1(t))}{c^2 \left(1 - \frac{|v_1(t)|^2}{c^2}\right)^2}$$

$$f(1, t, 1) = 0 \quad \text{for } |t| > 1$$

It is easy to verify that  $\Gamma_{XY} = \langle P, \mathcal{U}, m, s, f, c \rangle$  is a S.R.P.M., and consequently, so is  $\langle \Phi(\Gamma_{XY}), c' \rangle$ . Thus there is a positive number  $\delta$  such that

$$\Phi_1(m(1)) = \delta$$

We note next that at  $t=0$ :

$$s_1(0) = 0$$

$$(Ds_1)(0) = Y$$

$$(D^2 s_1)(0) = Z$$

and

$$f(1, 0, 1) = X$$

We thus have from (37) for  $t=0$

$$\delta \left[ \frac{d^2 q_1}{d\tau^2} (H_1(0)) \right] = \Phi_3(X, Y)$$

and thus from (45)

$$\Phi_3(X, Y) = \frac{X c'^2}{c^2 \lambda^2} \left[ \frac{d^2 q_1}{d\tau^2} (0) Q \right]_{1, \dots, n};$$

and hence from (36)

$$(46) \quad \varphi_3(X, Y) = \frac{\gamma c'^2}{c^2 \lambda^2} \left[ \left\langle X, \frac{X \cdot Y}{c^2} \right\rangle \alpha \right]_{1, \dots, n}.$$

In view of (1), (46) also holds for  $X=0$ .

Now let  $x$  be any positive real number. Then we set:

$$\delta = \langle 1, \dots, 1 \rangle$$

$$P = \{1\}$$

$$m(1) = x$$

$$\mathcal{I}(1) = (-c, c),$$

for  $t$  in  $\mathcal{I}(1)$

$$s_1(t) = \frac{1}{2} t^2 \delta$$

$$f(1, t, 1) = \frac{x \delta}{1 - \frac{|v_1(t)|^2}{c^2}} + \frac{x v_1(t) (\delta \cdot v_1(t))}{c^2 \left( 1 - \frac{|v_1(t)|^2}{c^2} \right)^2}$$

$$f(1, t, i) = 0 \quad \text{for } i > 1.$$

We easily verify that  $\Gamma_x = \langle P, \mathcal{I}, m, s, f, c \rangle$  is a S.R.P.M. such that for all  $t$  in  $\mathcal{I}(1)$ ,  $\sum_{i=1}^{\infty} f^{\text{rel}}(1, t, i) \neq 0$ . Furthermore, we infer from (36), (37), and (45) that for every  $t$  in  $\mathcal{I}(1)$

$$(47) \quad \sum_{i=1}^{\infty} f^{\text{rel}'}(1, H_1(t), i) = \frac{\varphi_1(x)}{x} \frac{c'^2}{c^2 \lambda^2} \sum_{i=1}^{\infty} f^{\text{rel}}(1, t, i) \alpha;$$

and hence from (46)

$$(48) \quad \varphi_1(x) = \gamma x.$$

Our theorem now follows from (19), (33), (34), (46), and (48).

**Remark 3.** We want to emphasize the physically reasonable nature of the hypothesis of the theorem just proved. We have assumed that systems of relativistic mechanics are carried by our transformations into systems of relativistic mechanics and that light lines are not carried into particle paths. No assumptions concerning the continuity of either  $\varphi_1$ ,  $\varphi_2$ , or  $\varphi_3$

have been made. Our assumption that  $\varphi_2$  is 1-1 may be justified physically by the argument that any two space-time positions of a particle distinct with respect to one observer must be distinct with respect to every observer.

The standard presentations of the special theory of relativity vary a good deal in their "derivations" of the Lorentz transformations. Almost without exception, however, the assumptions underlying these derivations are not clearly and completely stated. For the physicist who wants to begin with a set of axioms for relativistic particle mechanics with respect to a fixed coordinate system, our Theorem 2 provides a rigorous approach to the derivation of the Lorentz transformations. The transformations we obtain in Theorem 2 are, of course, more general than the Lorentz transformations, but it is obvious how the hypothesis of Theorem 2 may be strengthened so as to obtain just the ordinary Lorentz transformations.

Theorem 2 is also pertinent to discussions of the relativity of size, (see, for example, Hoffman [4]), since the determination of  $\varphi_1$ ,  $\varphi_2$ , and  $\varphi_3$  tells us exactly how the system of units of measurement may be changed in passing from one inertial frame of reference to another.

It is interesting to note that the set of transformations admissible (i.e., satisfying the hypothesis of Theorem 2) in relativistic particle mechanics differs sharply from the set of those admissible (see the hypothesis of Theorem 3 of McKinsey-Suppes [8]) in classical particle mechanics: in the latter case, but not in the former, admissible transformations can change the unit of distance differently along different coordinates (with correspondingly different changes in the unit of force). Thus, although classical mechanics can in a certain sense be regarded as a limiting case of relativistic mechanics, the set of transformations admissible in classical mechanics is in no sense a limit of the set of transformations admissible in relativistic mechanics.

5. Algebraic Structure of the Set of Admissible Transformations.

Let  $\Phi$  be an eligible transformation which satisfies the hypothesis of Theorem 2 with respect to the positive real numbers  $c$  and  $c'$ . We then call the ordered triple  $\langle \Phi, c, c' \rangle$  an admissible triple; and, corresponding to the informal usage at the end of the previous section, we call an eligible transformation an admissible transformation if it is the first element of some admissible triple. Since the set of admissible transformations is not a group under the obvious operation of composition, it is natural to ask what is its algebraic structure. We shall show that the structure of the set of admissible triples is that of a Brandt groupoid (formally defined below). Roughly speaking, the main difference between Brandt groupoids and groups is that a Brandt groupoid is not assumed to be closed under the binary operation corresponding to the group operation. Consequently, a Brandt groupoid may contain many identity elements, that is, many elements  $e$  such that  $x * e = x = e * x$  whenever  $x$ ,  $x * e$ , and  $e * x$  are in the groupoid. If there is an  $e$  in the groupoid such that for all  $x$  in the groupoid  $e * x = x = e * x$ , then the groupoid is also a group. For this reason, we introduced the notion of an admissible triple: the admissible transformation which carries every S.R.P.M. into itself is an identity element whose composition with every admissible transformation is defined; consequently, the set of admissible transformations is neither a group nor a Brandt groupoid.

The notion of a Brandt groupoid was first defined in Brandt [1]; we use the formal definition given in Jónsson-Tarski [5].

Definition 4. An algebraic system  $\mathcal{G} = \langle G, *, J, {}^{-1} \rangle$  (where  $*$  is an operation on a subset of  $U \times U$  to  $U$ ,  $J$  is a subset of  $U$  and  ${}^{-1}$  is an operation on  $U$  to  $U$ ) is called a BRANDT GROUPOID if and only if the following conditions are satisfied:



- (i) For  $x, y, z$  in  $G$ , if  $x * y \in G$  and  $y * z \in G$ , then  $(x * y) * z \in G$  and  $(x * y) * z = x * (y * z)$ .
- (ii) For  $x, y, z$  in  $G$ , if  $x * y \in G$  and  $x * y = x * z$ , then  $y = z$ .
- (iii) For  $x, y, z$  in  $G$ , if  $x * z \in G$  and  $x * z = y * z$ , then  $x = y$ .
- (iv)  $x * x = x$  for every  $x$  in  $J$ .
- (v)  $x^{-1} * x \in J$  and  $x * x^{-1} \in J$  for every  $x$  in  $G$ .
- (vi) For  $x, z$  in  $J$ , there exists a  $y$  in  $G$  such that  $x * y \in G$  and  $y * z \in G$ .

Rather than deal directly with admissible triples, it is somewhat simpler to use the following representation. From Theorem 2 we conclude that to each admissible triple there corresponds a unique ordered sextuple  $\langle A, B, \gamma, \lambda, c, c' \rangle$ , where  $B$  is an  $(n+1)$ -dimensional vector,  $\gamma, \lambda, c$ , and  $c'$  are positive real numbers, and  $A$  is a generalized Lorentz matrix (of order  $n+1$ ) with respect to  $\langle c, c', \lambda \rangle$ . Such an ordered sextuple  $\langle A, B, \gamma, \lambda, c, c' \rangle$  we shall call a carrier. From Theorem 1, together with Theorem 2, it then follows that there is a 1-1 correspondence between the set of carriers and the set of admissible triples.

We say that the carrier  $\langle A', B', \gamma', \lambda', c_1, c_2 \rangle$  is left-conformable to the carrier  $\langle A, B, \gamma, \lambda, c_3, c_4 \rangle$  if and only if  $c_1 = c_4$ . By the conformable subset  $\mathcal{S}$  of  $K \times K$  we mean the set of ordered pairs of elements of  $K$  such that the first element is left-conformable to the second.

We now define what we call the carrier system.

Definition 5. By THE CARRIER SYSTEM we mean the ordered quadruple  $\mathcal{K} = \langle K, *, J, {}^{-1} \rangle$ , where:

- (i)  $K$  is the set of all carriers;
- (ii)  $*$  is the operation on  $\mathcal{S}$  to  $K$  such that if the carrier  $\langle A', B', \gamma', \lambda', c_1, c_2 \rangle$  is left-conformable to the carrier  $\langle A, B, \gamma, \lambda, c_3, c_4 \rangle$  then

$$\begin{aligned} \langle A', B', \gamma', \lambda', c_1, c_2 \rangle * \langle A, B, \gamma, \lambda, c_3, c_4 \rangle \\ = \langle AA', BA' + B', \gamma\gamma', \lambda\lambda', c_3, c_2 \rangle \quad ; \end{aligned}$$

(iii)  $J$  is the set of carriers of the form  $\langle I, 0, 1, 1, c, c \rangle$ , where  $I$  is the identity matrix of order  $n+1$ ; and

(iv)  $^{-1}$  is the operation on  $K$  to  $K$  such that if  $\langle A, B, \gamma, \lambda, c, c' \rangle \in K$  then

$$\langle A, B, \gamma, \lambda, c, c' \rangle^{-1} = \langle A^{-1}, -BA^{-1}, 1/\gamma, 1/\lambda, c', c \rangle .$$

We have then the following theorem, the proof of which we omit.

Theorem 3. The carrier system is a Brandt groupoid.

We remark first that the operation  $*$  of the carrier system corresponds to the composition of admissible triples, i.e., if  $\langle \Phi, c, c' \rangle$  corresponds to  $\langle A, B, \gamma, \lambda, c, c' \rangle$  and  $\langle \Psi, c', c'' \rangle$  corresponds to  $\langle A', B', \gamma', \lambda', c', c'' \rangle$ , then  $\langle AA', BA' + B', \gamma\gamma', \lambda\lambda', c, c'' \rangle$  corresponds to  $\langle \Theta, c, c'' \rangle$ , where  $\langle \Theta, c, c'' \rangle$  is the admissible triple such that for any S.R.P.M.  $\Gamma_c$

$$\langle \Psi(\langle \Phi(\Gamma_c), c' \rangle), c'' \rangle = \langle \Theta(\Gamma_c), c'' \rangle .$$

Similarly, the inverse operation  $^{-1}$  of the carrier system corresponds to the natural inverse operation on admissible triples: i.e., if  $\langle \Phi, c, c' \rangle$  corresponds to  $\langle A, B, \gamma, \lambda, c, c' \rangle$  and  $\langle \Psi, c', c \rangle$  corresponds to  $\langle A^{-1}, -BA^{-1}, 1/\gamma, 1/\lambda, c', c \rangle$ , then for any S.R.P.M.  $\Gamma_c$

$$\langle \Psi(\langle \Phi(\Gamma_c), c' \rangle), c \rangle = \Gamma_c .$$

It thus follows as a corollary to Theorem 3 that the set of admissible triples is a Brandt groupoid under the natural operations of composition and formation of inverses.

It is natural to ask how the hypothesis of Theorem 2 may be strengthened so that the set of eligible transformations satisfying it form a group. We state without proof some results concerning this question.

Theorem 4. Let  $\Phi = \langle \varphi_1, \varphi_2, \varphi_3 \rangle$  be an eligible transformation which carries every system of relativistic particle mechanics into a system of relativistic particle mechanics. Then there exist positive real numbers  $\delta$ ,  $\gamma$ ,  $\lambda$ , and  $\rho$ , an  $(n+1)$ -dimensional vector  $B$ , an orthogonal matrix  $\xi$  of order  $n$  and a matrix  $A$  of order  $n+1$  such that

$$\delta^2 = 1$$

$$A = \lambda \begin{pmatrix} \delta & 0 \\ 0 & \rho \end{pmatrix} \begin{pmatrix} \xi & 0 \\ 0 & \delta \end{pmatrix},$$

and for any vectors  $z_1$  and  $z_2$  in  $E_n$ , any  $x$  in  $R$  and  $y$  in  $R^+$ ,

$$\varphi_1(y) = \gamma y$$

$$\varphi_2(z_1, x) = \langle z_1, x \rangle A + B,$$

$$\varphi_3(z_1, z_2) = \frac{\gamma \rho^2 z_1 \xi}{\lambda}.$$

The interpretation of  $\delta$ ,  $\gamma$ ,  $\lambda$ ,  $B$  and  $\xi$  is the same as that stated in Remark 1. The number  $\rho$  is the ratio  $c/c'$  of the absolute values of the old and new velocities of light. The matrix  $A$  is a generalized Lorentz matrix with  $U=0$ , which intuitively means that the old and new spatial frames of reference are at rest with respect to each other. The fact that the hypothesis of Theorem 4 thus excludes the possibility of transforming from one inertial frame of reference to another moving with respect to it is sufficient reason to regard this hypothesis as unnecessarily strong from the point of view of our intended physical interpretation. On the other hand, it is, of course, clear that the set of transformations satisfying this hypothesis constitute a group under the obvious operations.

Stanford University

REFERENCES

- [1] H. Brandt, "Über eine Verallgemeinerung des Gruppenbegriffes", Mathematische Annalen, vol. 96 (1926-1927), pp. 360-366.
- [2] P. Destouches-Février, La Structure des Théories Physiques, Paris: 1951.
- [3] H. Hermes, "Eine Axiomatisierung des allgemeinen Mechanik", Heft 3, Forschungen zur Logik und zur Grundlegung der Exakten Wissenschaften, Neue Folge, Leipzig: 1938.
- [4] B. Hoffmann, "The relativity of size", Physical Review, vol. 89 (1953), pp. 49-52.
- [5] B. Jónsson and A. Tarski, "Boolean algebras with operators, Part II", American Journal of Mathematics, vol. 74 (1952), pp. 127-162.
- [6] L. A. MacColl, "Geometric aspects of relativistic dynamics," Transactions of the American Mathematical Society, vol. 46 (1939), pp. 328-347.
- [7] J. C. C. McKinsey, A. C. Sugar, and P. Suppes, "Axiomatic foundations of classical particle mechanics", Journal of Rational Mechanics and Analysis, vol. 2 (1953), pp. .
- [8] J. C. C. McKinsey and P. Suppes, "Transformations of systems of classical particle mechanics," Journal of Rational Mechanics and Analysis, vol. 2 (1953), pp.
- [9] J. C. C. McKinsey and P. Suppes, "Philosophy and the axiomatic foundations of physics", Proceedings of the XIth International Congress of Philosophy, vol. \_\_\_\_ (1953).
- [10] K. Menger, "Are variables necessary in calculus", American Mathematical Monthly, vol. 56 (1949), pp. 609-620.
- [11] H. Reichenbach, Axiomatik der relativistischen Raum-Zeit-Lehre, Braunschweig: 1924.
- [12] A. A. Robb, Geometry of Time and Space, Cambridge: 1936.
- [13] K. Schnell, Eine Topologie der Zeit in logistischer Darstellung, Dissertation, Munster: 1937.
- [14] A. G. Walker, "Foundations of relativity I, II", Proceedings of the Royal Society of Edinburgh, Section A, vol. 62 (1948), pp. 319-335.

# FOOTNOTES

<sup>1/</sup>(page 1) We are grateful to Professor J. C. C. McKinsey for a large number of helpful suggestions and criticisms. This work was supported in part by a grant from the Office of Naval Research.

<sup>2/</sup>(page 6) Paulette Destouches-Février (in [2], pp. 5-6) advocates the use of a three-valued logic to describe the creation and annihilation of elementary particles. Actually, the situation is easily handled by the simple device of introducing the function  $\mathcal{I}$  defined on  $P$  instead of a fixed interval  $T$  for the whole system. Indeed, to our mind, her drastic proposal cannot be taken seriously until we know a great deal more about the mathematics which goes with a multi-valued logic. Even if such a body of mathematics existed (as it does not -- we do not have even the general outlines of elementary set theory in three-valued logic), it would be reasonable to adopt such a proposal only after every feasible alternative in standard mathematics had been explored.

<sup>3/</sup>(page 8) The intuitive interpretation of  $E_{n+1}$  is as the space-time manifold of special relativity with the  $(n+1)^{st}$  coordinate representing the time coordinate. Thus, if  $\langle Z, x \rangle$  is a point of  $E_{n+1}$ , then under the intended interpretation, the  $n$ -dimensional vector  $Z$  gives the spatial coordinates of the point and  $x$  its time coordinate.

<sup>4/</sup>(page 15) Readers familiar with the standard treatments of relativistic mechanics will note that (in the interests of rigor and explicitness) we have replaced " $t$ " by " $h_p(t)$ ".

<sup>5/</sup>(page 18) The statement of Theorem 2 would be made more symmetrical to Theorem 1 if  $\Phi_2$  were replaced by two functions  $\Phi'$  and  $\Phi''$  such that

$$\Phi'(Z, x) = [\Phi_2(Z, x)]_{n+1}$$

and

$$\Phi''(Z, x) = [\Phi_2(Z, x)]_{1, \dots, n}$$

This procedure was followed in McKinsey and Suppes [8] for classical mechanics, but in relativistic mechanics, it is natural to introduce the single transformation  $\Phi_2$  for the space-time manifold.